

RECEIVED

MAR 28 2003  
TECH CENTER 1800/2800

1600

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/597,796C

DATE: 03/20/2003 P.6

TIME: 16:10:20

Input Set : A:\-90-5.app

Output Set: N:\CRF4\03202003\I597796C.raw

```

3 <110> APPLICANT: Skeiky, Yasir
4     Reed, Steven
5     Alderson, Mark
6     Corixa Corporation
8 <120> TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
10 <130> FILE REFERENCE: 014058-009050US
12 <140> CURRENT APPLICATION NUMBER: US 09/597,796C
13 <141> CURRENT FILING DATE: 2000-06-20
15 <150> PRIOR APPLICATION NUMBER: US 09/056,556
16 <151> PRIOR FILING DATE: 1998-04-07
18 <150> PRIOR APPLICATION NUMBER: US 09/223,040
19 <151> PRIOR FILING DATE: 1998-12-30
21 <150> PRIOR APPLICATION NUMBER: WO PCT/US99/07717
22 <151> PRIOR FILING DATE: 1999-04-07
24 <150> PRIOR APPLICATION NUMBER: US 09/287,849
25 <151> PRIOR FILING DATE: 1999-04-07
27 <150> PRIOR APPLICATION NUMBER: US 60/158,338
28 <151> PRIOR FILING DATE: 1999-10-07
30 <150> PRIOR APPLICATION NUMBER: US 60/158,425
31 <151> PRIOR FILING DATE: 1999-10-07
33 <160> NUMBER OF SEQ ID NOS: 30
35 <170> SOFTWARE: PatentIn Ver. 2.1
37 <210> SEQ ID NO: 1
38 <211> LENGTH: 588
39 <212> TYPE: DNA
40 <213> ORGANISM: Mycobacterium tuberculosis
42 <220> FEATURE:
43 <223> OTHER INFORMATION: Ra35, N-terminus of MTB32A (TbRa35FL)
45 <220> FEATURE:
46 <221> NAME/KEY: CDS
47 <222> LOCATION: (1)..(588)
48 <223> OTHER INFORMATION: Ra35
50 <400> SEQUENCE: 1
51 gccccgccgg ccttgctcgca ggaccgggttc gccgacttcc ccgcgctgcc cctcgacccg 60
52 tccgcgatgg tcgcccaagt ggggccacag gtggtcaaca tcaacaccaa actggggtac 120
53 aacaacgccg tgggcgccgg gaccggcatc gtcatcgatc ccaacggtgt cgtgctgacc 180
54 aacaaccacg tgatcgccgg cgccaccgac atcaatgctg tcagcgctcg ctccggccaa 240
55 acctacggcg tcgatgtggt cgggtatgac cgcacccagg atgtcgcggt gctgcagctg 300
56 cgcggtgccg gtggcctacc atcggcgggc atcggtgggc gcgtcgcggt tggtgagccc 360
57 gtcgtcgcca tgggcaacag cggtgggcag ggcggaacgc cccgtgcggt gcctggcagg 420
58 ttggtcgcgc tcggccaaac cgtgcaggcg tcggattcgc tgaccggtgc cgaagagaca 480
59 ttgaacgggt tgatccagtt cgatgccgcg atccagcccg gtgattcggg cgggcccgtc 540
60 gtcaacggcc taggacaggt ggtcggtatg aacacggccg cgtcctag 588

```

## RAW SEQUENCE LISTING

DATE: 03/20/2003

PATENT APPLICATION: US/09/597,796C

TIME: 16:10:20

Input Set : A:\-90-5.app

Output Set: N:\CRF4\03202003\I597796C.raw

```

63 <210> SEQ ID NO: 2
64 <211> LENGTH: 195
65 <212> TYPE: PRT
66 <213> ORGANISM: Mycobacterium tuberculosis
68 <220> FEATURE:
69 <223> OTHER INFORMATION: Ra35, N-terminus of MTB32A (TbRa35FL)
71 <400> SEQUENCE: 2
72 Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
73   1           5           10           15
74 Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Gly Pro Gln Val Val
75           20           25           30
76 Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
77           35           40           45
78 Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
79           50           55           60
80 Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln
81   65           70           75           80
82 Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
83           85           90           95
84 Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
85           100          105          110
86 Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
87           115          120          125
88 Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
89           130          135          140
90 Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
91  145          150          155          160
92 Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser
93           165          170          175
94 Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
95           180          185          190
96 Ala Ala Ser
97           195
100 <210> SEQ ID NO: 3
101 <211> LENGTH: 1872
102 <212> TYPE: DNA
103 <213> ORGANISM: Mycobacterium tuberculosis
105 <220> FEATURE:
106 <223> OTHER INFORMATION: MTB32A (TbRa35FL) cDNA
108 <220> FEATURE:
109 <221> NAME/KEY: modified_base
110 <222> LOCATION: (1460)
111 <223> OTHER INFORMATION: n = g, a, c or t
113 <220> FEATURE:
114 <221> NAME/KEY: modified_base
115 <222> LOCATION: (1854)
116 <223> OTHER INFORMATION: n = g, a, c or t
118 <400> SEQUENCE: 3
119 gactacgttg gtgtagaaaa atcctgccgc ccggaccctt aaggctggga caatttctga 60

```

## RAW SEQUENCE LISTING

DATE: 03/20/2003

PATENT APPLICATION: US/09/597,796C

TIME: 16:10:20

Input Set : A:\-90-5.app

Output Set: N:\CRF4\03202003\I597796C.raw

```

120 tagctacccc gacacaggag gttacgggat gagcaattcg cgccgcccgt cactcagggtg 120
121 gtcattggttg ctgagcgtgc tggctgccgt cgggctgggc ctggccacgg cgccggccca 180
122 ggcggccccc cgcccttgt cgcaggaccg gtccgcccac tccccgcgc tgcccctcga 240
123 cccgtccgcg atggtcgccc aagtggcgcc acagggtggtc aacatcaaca ccaaactggg 300
124 ctacaacaac gccgtgggcg ccgggaccgg catcgtcatc gatccaacg gtgtcgtgct 360
125 gaccaacaac cacgtgatcg cgggcgccac cgacatcaat gcgttcagcg tcggctccgg 420
126 ccaaacctac ggctcgatg tggtcgggta tgaccgcacc caggatgtcg cggtgctgca 480
127 gctgcgcggt gccgtggcc tgccgtcggc ggcatcggt ggccggtcg cggttggtga 540
128 gcccgtcgtc gcgatgggca acagcgggtg gcaggcgga acgccccgtg cggtgccctg 600
129 cagggtggtc gcgtcggcc aaaccgtgca ggctcggtg tcgctgaccg gtgccgaaga 660
130 gacattgaac gggttgatcc agttcgatgc cgcaatccag cccggtgatt cgggcggggc 720
131 cgtcgtcaac ggctaggac aggtggtcgg tatgaacacg gccgcgtccg ataacttcca 780
132 gctgtccag ggtggcgagg gattcgccat tccgatcggg caggcgatgg cgatcgcggg 840
133 ccaaattccga tcgggtgggg ggtcacccac cgttcatac ggccctaccg ccttcctcgg 900
134 cttgggtgtt gtcgacaaca acggcaacgg cgcacgagtc caacgcgtgg tcggaagcgc 960
135 tccggcgcca agtctcgcca tctccaccgg cgacgtgat acccggtcg acggcgctcc 1020
136 gatcaactcg gccaccgcca tggcggacgc gcttaacggg catcatcccg gtgacgtcat 1080
137 ctccgtgaac tggcaaacca agtcgggagg cagcgtaca ggaacgtga cattggccga 1140
138 gggacccccg gcctgatttg tcgcggtac caccgcggc cggccaatt ggattggcgc 1200
139 cagccgtgat tgccgcgtga gccccgagt tccgtctccc gtgcgcgtgg cattgtggaa 1260
140 gcaatgaacg aggagaaca cagcgttgag caccctccc tgcagggcag ttacgtcgaa 1320
141 ggcggtgtgg tcgagcatcc ggatgccaag gacttcggca gcgcgcggc cctgccgcc 1380
142 gatccgacct ggtttaagca cgccgtcttc tacgaggtgc tggtcggggc gttcttcgac 1440
W--> 143 gccagcgcgg acggttccgn cgatctgagt ggactcatcg atcgccctga ctacctgcag 1500
144 tggcttgcca tcgactgcat ctgttgccgc cgttcctacg actcaccgct gcgcgacggc 1560
145 ggttacgaca ttcgcgactt ctacaagggt ctgcccgaat tcggcaccgt cgacgatttc 1620
146 gtcgccctgg tcgacaccgc tcaccggcga ggtatccgca tcatcaccga cctggtgatg 1680
147 aatcacacct cggagtcgca cccctggttt caggagtcgc gccgcgacc agacggaccg 1740
148 tacggtgact attacgtgtg gagcgacacc agcgagcgct acaccgacgc ccggtatc 1800
149 ttcgtcgaca ccgaagagtc gaactggtca ttcgatcctg tccgcccgaca gtttctactg 1860
150 gcaccgattc tt 1872
153 <210> SEQ ID NO: 4
154 <211> LENGTH: 355
155 <212> TYPE: PRT
156 <213> ORGANISM: Mycobacterium tuberculosis
158 <220> FEATURE:
159 <223> OTHER INFORMATION: MTB32A (TbRa35FL) protein
161 <400> SEQUENCE: 4
162 Met Ser Asn Ser Arg Arg Arg Ser Leu Arg Trp Ser Trp Leu Leu Ser
163 1 5 10 15
164 Val Leu Ala Ala Val Gly Leu Gly Leu Ala Thr Ala Pro Ala Gln Ala
165 20 25 30
166 Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
167 35 40 45
168 Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Ala Pro Gln Val Val
169 50 55 60
170 Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
171 65 70 75 80
172 Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val

```

## RAW SEQUENCE LISTING

DATE: 03/20/2003

PATENT APPLICATION: US/09/597,796C

TIME: 16:10:20

Input Set : A:\-90-5.app

Output Set: N:\CRF4\03202003\I597796C.raw

```

173           85           90           95
174 Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln
175           100           105           110
176 Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
177           115           120           125
178 Val Leu Gln Leu Arg Gly Ala Gly Gly Leu Pro Ser Ala Ala Ile Gly
179           130           135           140
180 Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
181 145           150           155           160
182 Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
183           165           170           175
184 Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
185           180           185           190
186 Leu Asn Gly Leu Ile Gln Phe Asp Ala Ala Ile Gln Pro Gly Asp Ser
187           195           200           205
188 Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
189           210           215           220
190 Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe Ala
191 225           230           235           240
192 Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser Gly
193           245           250           255
194 Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu
195           260           265           270
196 Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val Val
197           275           280           285
198 Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val Ile
199           290           295           300
200 Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala Asp
201 305           310           315           320
202 Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp Gln
203           325           330           335
204 Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly
205           340           345           350
206 Pro Pro Ala
207           355
210 <210> SEQ ID NO: 5
211 <211> LENGTH: 447
212 <212> TYPE: DNA
213 <213> ORGANISM: Mycobacterium tuberculosis
215 <220> FEATURE:
216 <223> OTHER INFORMATION: MTBRa12 C-terminus of MTB32A (Ra35FL)
218 <400> SEQUENCE: 5
219 cggatatgaac acggccgcgt ccgataactt ccagctgtcc caggggtgggc agggattcgc 60
220 cattccgatc gggcaggcga tggcgatcgc gggccagatc cgatcgggtg gggggtcacc 120
221 caccgttcac atcgggccta ccgccttcct cggttgggt gttgtcgaca acaacggcaa 180
222 cggcgacaga gtccaacgcg tggtcgggag cgctccggcg gcaagtctcg gcatctccac 240
223 cggcgacgtg atcaccgcgg tcgacggcgc tccgatcaac tcggccaccg cgatggcgga 300
224 cgcgcttaac gggcatcatc ccggtgacgt catctcggtg aactggcaaa ccaagtcggg 360
225 cggcacgcgt acaggaacg tgacattggc cgagggaccc ccggcctgat ttcgtcgygg 420

```

## RAW SEQUENCE LISTING

DATE: 03/20/2003

PATENT APPLICATION: US/09/597,796C

TIME: 16:10:20

Input Set : A:\-90-5.app

Output Set: N:\CRF4\03202003\I597796C.raw

```

226 ataccacccg ccggccggcc aattgga 447
229 <210> SEQ ID NO: 6
230 <211> LENGTH: 132
231 <212> TYPE: PRT
232 <213> ORGANISM: Mycobacterium tuberculosis
234 <220> FEATURE:
235 <223> OTHER INFORMATION: MTBRa12 C-terminus of MTB32A (Ra35FL)
237 <400> SEQUENCE: 6
238 Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
239 1 5 10 15
240 Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser
241 20 25 30
242 Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly
243 35 40 45
244 Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val
245 50 55 60
246 Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val
247 65 70 75 80
248 Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala
249 85 90 95
250 Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp
251 100 105 110
252 Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu
253 115 120 125
254 Gly Pro Pro Ala
255 130
258 <210> SEQ ID NO: 7
259 <211> LENGTH: 3058
260 <212> TYPE: DNA
261 <213> ORGANISM: Mycobacterium tuberculosis
263 <220> FEATURE:
264 <223> OTHER INFORMATION: MTB39 (TbH9) cDNA full-length
266 <400> SEQUENCE: 7
267 gatcgtaccc gtgcgagtgc tcgggcccgtt tgaggatgga gtgcacgtgt ctttcgtgat 60
268 ggcataccca gagatgttgg cggcggcggc tgacaccctg cagagcatcg gtgctaccac 120
269 tgtggctagc aatgccgctg cggcggcccc gacgactggg gtggtgcccc ccgctgccga 180
270 tgagggtgctg gcgctgactg cggcgcactt cgccgcacat gcggcgatgt atcagtccgt 240
271 gagcgctcgg gctgctgcga ttcattgacca gttcgtggcc acccttgcca gcagcgccag 300
272 ctctgatgctg gccactgaag tcgccaatgc ggcggcggcc agctaagcca ggaacagtcg 360
273 gcacgagaaa ccacgagaaa tagggacacg taatggtgga tttcggggcg ttaccaccgg 420
274 agatcaactc cgcgaggatg tacgccggcc cgggttcggc ctgctgtgtg gccgcggctc 480
275 agatgtggga cagcgtggcg agtgacctgt tttcggccgc gtcggcggtt cagtcggtgg 540
276 tctggggctc gacggtgggg tcgtggatag gttcgtcggc gggctctgat gtggcggcgg 600
277 cctcgccgta tgtggcgtgg atgagcgtca ccgcggggca ggccgagctg accgccgcc 660
278 aggtccgggt tgctgcggcg gcctacgaga cggcgatagg gctgacgggt ccccgccgg 720
279 tgatgccga gaacctgtc gaactgatga ttctgatagc gaccaacctc ttggggcaaa 780
280 acaccccggc gatcgcggtc aacgaggccg aatacggcga gatgtgggcc caagacgccg 840
281 ccgcgatgtt tggctacgcc gcggcgacgg cgacggcgac ggcgacgttg ctgccgttcg 900
282 aggaggcgcc ggagatgacc agcgcggggtg ggctcctcga gcaggccgcc gcggtcgagg 960

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/597,796C

DATE: 03/20/2003  
TIME: 16:10:21

Input Set : A:\-90-5.app  
Output Set: N:\CRF4\03202003\I597796C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 1460,1854  
Seq#:11; N Pos. 30,33,2270  
Seq#:17; N Pos. 497,500,1136,1445,1487,1509,1515  
Seq#:25; N Pos. 767  
Seq#:26; Xaa Pos. 254

**VERIFICATION SUMMARY**

DATE: 03/20/2003

PATENT APPLICATION: US/09/597,796C

TIME: 16:10:21

Input Set : A:\-90-5.app

Output Set: N:\CRF4\03202003\I597796C.raw

L:143 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:1440  
M:341 Repeated in SeqNo=3  
L:546 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0  
M:341 Repeated in SeqNo=11  
L:801 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:480  
M:341 Repeated in SeqNo=17  
L:1100 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25 after pos.:720  
L:1149 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26 after pos.:240  
L:1204 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:27